 Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 Period \_\_\_\_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |
| --- |
| **3.1 Scatterplots and Correlation** |
| Why do we study relationships between two variables? |  |
| **Types of Variables** |
| Explanatory Variable |  |
| Response Variable |  |
| Examples |  |
| **Scatterplots** |
| What is a scatterplot? |  |
| Things to Remember When Making Scatterplots |  |
| *Problem 1 – Track and Field Day!*The table below shows data for 13 students in a statistics class. Each member of the class ran a 40-yard sprint and then did a long jump (with a running start). Make a scatterplot of the relationship between sprint time (in seconds) and long jump distance (in inches).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sprint Time (s) | 5.41 | 5.05 | 9.49 | 8.09 | 7.01 | 7.17 | 6.83 | 6.73 | 8.01 | 5.68 | 5.78 | 6.31 | 6.04 |
| Long Jump Distance (in)  | 171 | 184 | 48 | 151 | 90 | 65 | 94 | 78 | 71 | 130 | 173 | 143 | 141 |

**Texas Instruments TI-84 Plus Silver Edition Graphing Calculator in Black** |
| *Problem 2* – *Puppies and Snowshoes* Make a scatterplot of the relationship between weight of a puppy and snowshoe price.

|  |
| --- |
| **Weekly Data Collection** |
| **The weight of a growing puppy in New York** | **The retail price of snowshoes in Alaska** |
| 8 pounds | $32.45 |
| 8.5 | $32.95 |
| 9 | $33.45 |
| 9.6 | $34.00 |
| 10.1 | $34.50 |
| 10.7 | $35.10 |
| 11.5 | $35.63 |

 |
| How to Examine a Scatterplot |  |
| Describe the Track and Field Data | Describe the Puppy and Snowshoe Data |
| Does a strong association between 2 variables indicate a cause-and-effect relation? |  |
| How can we measure the strength and direction? |  |
| Characteristics of correlation |  |
| How is the correlation coefficient calculated? |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Texas Instruments TI-84 Plus Silver Edition Graphing Calculator in Black***Problem 3 – Speed and Mileage*Calculate the correlation coefficient for the following data set. Explain the results.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Speed  | 20 | 30 | 40 | 50 | 60 |
| Mileage | 24 | 28 | 30 | 28 | 24 |

 |
| Important Facts about Correlation |  |
| **3.2 Least-Squares Regression** |
| What is a regression line and what is it used for? |  |
| What is the general form of a regression equation? |  |
| How are the slope and y-intercept of a regression line interpreted? |  |
| *Problem 4 – Used Hondas*The following scatterplot shows the number of miles driven (in thousands) and advertised price (in thousands) for 11 used Honda CR-Vs from the 2002-2006 model years. The regression line shown on the scatterplot is Cost = 18773 – .08618(Miles Driven). (a) Interpret the slope and *y* intercept in context. (b) Predict the advertised price of a Honda CR-V that has 35,600 miles. |
| What is extrapolation and why is it dangerous? |  |
| The Least-squares Regression Line  |  |
| Finding the LSRL and the correlation coefficient using the calculator | **Texas Instruments TI-84 Plus Silver Edition Graphing Calculator in Black** |
| *Problem 5 – Track and Field Day Revisited* The table below shows data for 13 students in a statistics class. The table shows the height of each member of the class and the long jump (with a running start) distance.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Height (in) | 74 | 72 | 60 | 70 | 65 | 65 | 68 | 69 | 64 | 70 | 71 | 69 | 72 |
| Long Jump Distance (in)  | 171 | 184 | 48 | 151 | 90 | 65 | 94 | 78 | 71 | 130 | 173 | 143 | 141 |

1. Calculate the equation of the least-squares regression line using technology.
2. Interpret the slope in context. Does it make sense to interpret the *y* intercept?

(c) Identify and interpret the correlation coefficient. What is the relationship between the correlation coefficient and the slope of the regression line?  |
| **Residuals and Residual Plots** |
| What is a **residual**? How do you interpret a residual? |    |
| *Problem 6 – Height vs Long Jump*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Height (in) | 74 | 72 | 60 | 70 | 65 | 65 | 68 | 69 | 64 | 70 | 71 | 69 | 72 |
| Long Jump Distance (in)  | 171 | 184 | 48 | 151 | 90 | 65 | 94 | 78 | 71 | 130 | 173 | 143 | 141 |

1. Calculate and interpret the residual for the athlete who has a height of 64”.
2. Calculate and interpret the residual for the athlete who has a height of 74”.

  |
| Finding Residuals using the Calculator | **Texas Instruments TI-84 Plus Silver Edition Graphing Calculator in Black** |
| *Problem 6 (cont.) – Track and Field Day Revisited* 1. Find and interpret the residual for student who is 71” tall?
 |
| What is a residual plot? What is the purpose of a residual plot? |  |
| What two things do you look for in a residual plot? | 1)**http://ebooks.bfwpub.com/tps3e/figures/3_19.gif** |
|  | 2)  |
| Construct and interpret a residual plot for the Track-and-Field Day data.  | **Texas Instruments TI-84 Plus Silver Edition Graphing Calculator in Black** |
| Standard Deviation of the residuals (*s* ) |  |
| **The Coefficient of Determination, *r* 2**  |
| *Problem 7 – The Unknown Height*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Height (in) | 74 | 72 | 60 | 70 | 65 | 65 | 68 | 69 | 64 | 70 | 71 | 69 | 72 |
| Long Jump Distance (in)  | 171 | 184 | 48 | 151 | 90 | 65 | 94 | 78 | 71 | 130 | 173 | 143 | 141 |

Suppose that a student was absent on the day these data were collected. Predict this student’s long jump distance.  |
| What is *r* 2? How do you calculate and interpret *r* 2? |    |
| Compare and Contrast *s* and *r* 2 |  |
| *Problem 8 – Interpreting Computer Regression Output* A random sample of 15 high school students was selected from the U.S. CensusAtSchool database. The foot length (in centimeters) and height (in centimeters) of each student in the sample were recorded. Least-squares regression was performed on the data. A scatterplot with the regression line added, a residual plot, and some compute routput from the regression are shown below:1. What is the equation of the least-squares regression line that describes the relationship between foot length and height?
2. Interpret the slope of the regression line in context.
3. Find and interpret the correlation.
4. Is a line an appropriate model to use for these data? Explain.
 |
| How can we calculate the Least-Squares Regression line?  |  | **Example** |
| **Correlation and Regression Wisdom**  |
| Things to consider and correlation and regression |  |
| How do outliers influence regression? | **http://ebooks.bfwpub.com/tps3e/figures/3_23_big.gif** |